

Cloning and characterization of *mariner*-like elements in the soybean aphid, *Aphis glycines* Matsumura

O. Mittapalli^{1*}, L. Rivera-Vega¹, B. Bhandary¹,
M.A. Bautista¹, P. Mamidala¹, A.P. Michel¹,
R.H. Shukle² and M.A.R. Mian³

¹Department of Entomology, The Ohio State University / Ohio Agricultural and Research Development Center, 1680 Madison Avenue, Wooster, OH 44691, USA; ²USDA-ARS and Department of Entomology, Purdue University, West Lafayette, IN 47907, USA; ³USDA-ARS and Department of Horticulture and Crop Science, The Ohio State University, 1680 Madison Avenue, Wooster, OH 44691, USA

Abstract

Soybean aphid, *Aphis glycines* Matsumura (Hemiptera: Aphididae), is currently the most important insect pest of soybean (*Glycine max* (L.) Merr.) in the United States and causes significant economic damage worldwide, but little is known about the aphid at the molecular level. *Mariner*-like transposable elements (MLEs) are ubiquitous within the genomes of arthropods and various other invertebrates. In this study, we report the cloning of MLEs from the soybean aphid genome using degenerate PCR primers designed to amplify conserved regions of *mariner* transposases. Two of the ten sequenced clones (designated as *Agmar1* and *Agmar2*) contained partial but continuous open reading frames, which shared high levels of homology at the protein level with other *mariner* transposases from insects and other taxa. Phylogenetic analysis revealed *Agmar1* to group within the irritans subfamily of MLEs and *Agmar2* within the mellifera subfamily. Southern blot analysis and quantitative PCR analysis indicated a low copy number for *Agmar1*-like elements within the soybean aphid genome. These results suggest the presence of at least two different putative *mariner*-like transposases encoded by the soybean aphid genome. Both *Agmar1* and *Agmar2* could play influential roles in the architecture of the soybean aphid genome. Transposable elements are also thought to potentially mediate resistance in insects through changes in gene amplification and mutations in coding sequences. Finally, *Agmar1* and *Agmar2* may represent useful genetic tools and provide insights on *A. glycines* adaptation.

Keywords: *Agmar1*, *Agmar 2*, *mariner*-like element, soybean aphid, transposable element

(Accepted 4 April 2011)

*Authors for correspondence
Fax: +01 330-263-3686
E-mail: mittapalli.1@osu.edu

Introduction

The soybean aphid, *Aphis glycines* Matsumura (Hemiptera: Aphididae), is the most important insect pest of Soybean (*Glycine max* (L.) Merr.). Soybean aphids damage by feeding directly on plants and as a vector of various viruses (Wu *et al.*, 2004). The United States is the world's largest producer of soybean (91.4 million metric tons in the year 2009: FAO Statistical Database, 2010), but production in 2000 was threatened by the accidental introduction of the soybean aphid into North America (Ragsdale *et al.*, 2007). An estimated 80% of the North American soybean crop is considered at risk (Vennette & Ragsdale, 2004).

Insecticide use is the most effective and practical method for soybean aphid control, but the environmental hazards, particularly on natural enemies, has necessitated research in host-plant resistance to the soybean aphid. A number of soybean accessions (plant introductions/collections from other countries) with resistance to the soybean aphid have been identified (Hill *et al.*, 2004; Mensah *et al.*, 2005; Mian *et al.*, 2008a; Zhang *et al.*, 2009), and the two single dominant genes, *Rag1* (Hill *et al.*, 2004) and *Rag2* (Mian *et al.*, 2008b), have been mapped and transferred into commercial soybean cultivars. Commercial cultivars with *Rag1* were made available in 2010. The resistance by these genes in soybean is thought to operate via antibiosis, wherein aphids are killed as early nymphs. Before the widespread use of resistant cultivars, however, virulent soybean aphid populations have been recently discovered, termed as soybean aphid biotypes (Kim *et al.*, 2008; Hill *et al.*, 2010). At least three biotypes have been formally described, and the existence of others has been noted in field observations (A. Michel, personal communication). Biotype evolution has occurred through either genetic selection or recent changes at the genetic level. Answering this question depends on molecular resources which are currently lacking for this devastating insect pest.

Transposable elements belonging to the *mariner* family are thought to be the most widely distributed genetic elements in nature. These mobile elements are integral components in the genomes of various taxa, including bacteria, nematodes, insects, plants, fish and mammals (Hartl, 1989; Robertson, 1993; Robertson & Lampe, 1995a; Jarvik & Lark, 1998; Izsvák *et al.*, 1999; Leroy *et al.*, 2000; Schneider *et al.*, 2000; Liu *et al.*, 2007). *Mariner*-like elements (MLEs) are members of the *mariner*: Tc1 transposon superfamily. The first *mariner* element was identified and characterized by Jacobson *et al.* (1986) in the *white-peach* mutant of *Drosophila mauritiana* Tsacas and David (Diptera: Drosophilidae). These functional elements encode a single protein, a transposase, which allow them to spontaneously incorporate into genomes via a cut-and-paste mechanism. Other characteristic features of MLEs include the presence of a short inverted terminal repeat and the requirement of a DNA intermediate for transposition. In view of their unique characteristics at the genome level across insect species, MLEs have been proposed to serve as valuable research tools, such as vectors for germline transformation, isolation and mapping of genes and transposon-based sequencing (Cooley *et al.*, 1988; Strausbaugh *et al.*, 1990; Maruyama & Hartl, 1991).

In general, two forms of MLEs exist: active and inactive, also termed autonomous and nonautonomous elements (Medhora *et al.*, 1991). Inactive elements are simple derivatives of active forms with in-frame deletions that result in the loss of

essential sequences required for transposition but retain other functional features. Thus, they are able to transpose only in the presence of active elements, which are sometimes referred to as mutator factors (Medhora *et al.*, 1991). Many of the *mariner* elements are assumed to be inactive, and only four autonomous elements have been reported, including *Mos1* from *D. mauritiana* (Medhora *et al.*, 1991); *Famar1* from the European earwig, *Forficula auricularia* (Barry *et al.*, 2004); *Himar1* from the horn fly, *Haemotobia irritans* (Bryan *et al.*, 1987; Robertson & Lampe, 1995a); and *Mboumar-9* from the ant, *Messor bouvieri* (Muñoz-Lopez *et al.*, 2008).

Given that transposable elements could serve as genetic tools and/or potentially mediate resistance in insects through changes in gene amplification and mutations in coding sequences (Li *et al.*, 2007), we were interested in identifying MLEs in the soybean aphid genome. To our knowledge, this study would represent the first report of such functional elements and their implications in the soybean aphid genome.

Materials and methods

Experimental insect

The 'Illinois' (IL) colony samples were collected from a colony established in 2000 (Hill *et al.*, 2004). The 'Ohio' (OH) colony was established from field collected aphids in Wooster, OH in the summer of 2005 (Kim *et al.*, 2008). All colonies were established from a large number of randomly collected individuals from soybean. Individuals from each colony used in the current study were randomly collected from multiple soybean plants; only one aphid per plant was collected from randomly selected soybean plants to limit sampling of clones. The aphid colonies were maintained on seedlings of cultivar 'Williams 82' in growth chambers at temperatures between 22°C and 24°C with a photosynthetically active radiation of 330 $\mu\text{mol m}^{-2} \text{s}^{-1}$ for 15 h daily and 60–70% relative humidity (RH). DNA samples were stored at –80°C in the lead author's laboratory to serve as vouchers.

Homology-based cloning

Genomic DNA from 50 *A. glycines* individuals, including adults and nymphs, was isolated using the E.Z.N.A.[®] Tissue DNA Kit from Omega Biotek, Inc (Doraville, GA, USA) following the manufacturers protocol. DNA isolated using the same protocol from 50 adult orange blossom wheat midge, *Sitodiplosis mosellana* (known to carry MLEs in the genome: Mittapalli *et al.*, 2006), was used as the positive control in PCR reactions. DNA-negative controls were represented by water as the template in PCR reactions. Degenerate PCR primers designed to amplify regions of conserved amino acid residues in the middle of the *mariner* transposases as per Robertson (1993) with minor modifications were used (Mittapalli *et al.*, 2006). The sequences of the degenerate primers are: MAR-159F, 5'-TNACNDKNGAYGARAMNTGG-3' and MAR-249R, 5'-RYRTGNGGNSKNGCRTRTC-3', where D=A+G+T, K=G+T, M=A+C, N=A+C+G+T, R=A+G, S=C+G and Y=C+T. PCR cycling included: 94°C for 1 min, 55°C for 1 min and 72°C for 1 min for a total of 40 cycles, 72°C for 5 min, 4°C hold. Post amplification, 5 μl of each PCR reaction was analyzed on a 1% agarose gel. Four microliters of fresh PCR product was used to clone the amplified fragments directly

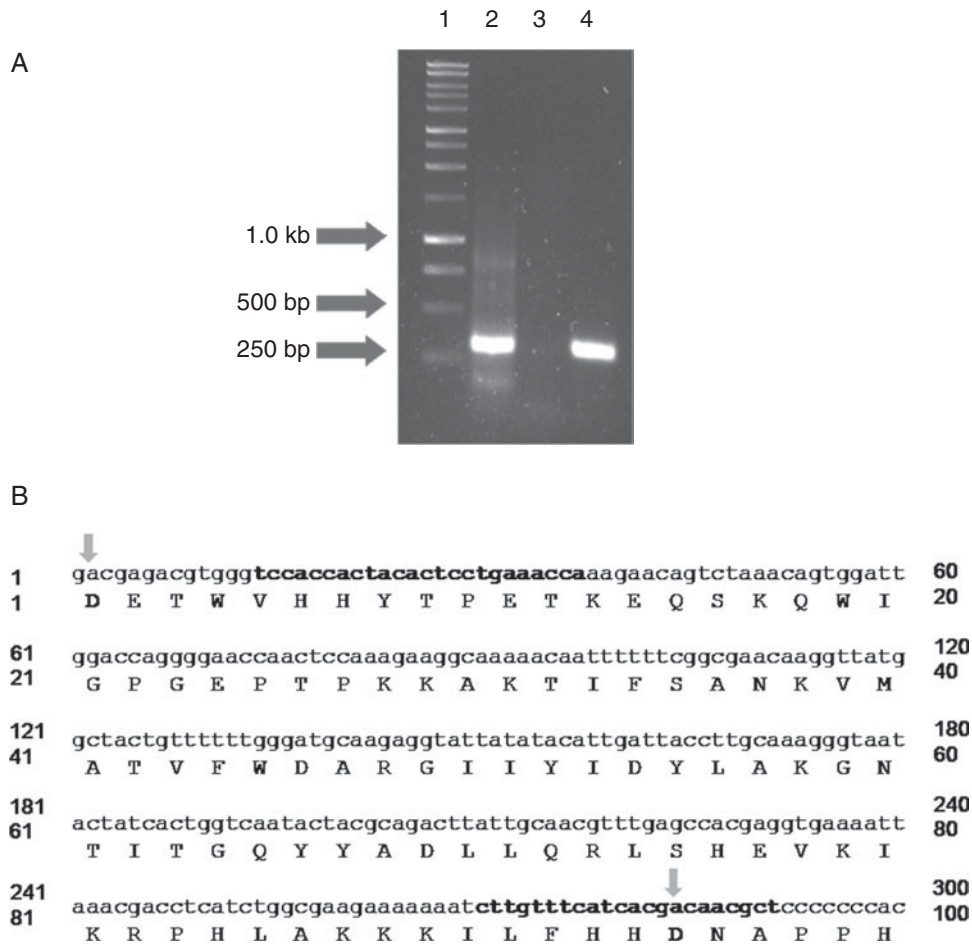


Fig. 1. *Mariner*-like elements (MLEs) in the soybean aphid, *Aphis glycines*. (a) PCR amplification of MLEs in the soybean aphid using degenerate PCR primers for conserved regions of *mariner* transposases. Lanes are as follows: 1, Promega 1-kb ladder; 2, PCR fragments amplified using *A. glycines* DNA; 3, DNA-negative control PCR; 4, DNA-positive control using Orange blossom wheat midge (*Sitodiplosis mosellana*) DNA. (b) Partial MLE sequences obtained from *A. glycines* (*Agmar1*) without in-frame terminations 'showing conserved domains of *mariner* transposases'. Highlighted with grey arrow-heads are two of the three aspartic acid (Asp, D) residues that constitute the D,D(34)D motif of the catalytic domain of *Agmar1*. Bold nucleotide sequences represent gene-specific primers for downstream analysis.

into the PCR[®] 4-TOPO[®] vector included in a TOPO TA cloning[®] for sequencing kit (Invitrogen, Carlsbad, CA, USA) following the manufacturers protocol.

Sequence analysis

Ten independent clones (*Agmar1*–10) were sequenced using M13 sequencing primers over both strands at the Molecular and Cellular Imaging Center (MCIC) located at the Ohio Agricultural and Research Development Center (OARDC). Sequence analyses, including similarity and annotation searches, were done using different BLAST programs (Altschul *et al.*, 1990) on the National Center for Biotechnology Information (Bethesda, MD, USA), web site (<http://www.ncbi.nlm.nih.gov/>). The deduced amino acid sequences for the soybean aphid MLEs were obtained using the 'Translate tool' (<http://au.expasy.org/tools/dna.html>). Multiple sequence alignment of the soybean aphid MLEs with MLE sequences from other insect species was performed using

Multiple Sequence Comparison by Log Expectation (MUSCLE: Edgar, 2004).

Phylogenetic analysis

In order to reveal the relationship of the soybean aphid MLE sequences to other *mariner* transposases, a phylogenetic analysis was performed. A phylogram was constructed using the deduced amino acid sequences of the soybean aphid MLEs and different *mariner* transposase subfamilies, including mauritiana, irritans, capitata, mellifera and cecropia. The peptide sequences were aligned with the ClustalX program, version 8.1 with 11 updates (Thompson *et al.*, 1997). The phylogram was calculated by the distance/neighbor-joining method using the software package PAUP* 4.0b10 (Swofford, 2002). The polypeptide sequence of a *Tc1* transposable element transposase from *Caenorhabditis elegans* (accession no. X01005) represented the outgroup. Bootstrap values for the branches were obtained with 1000 replications.

A

<i>Agmar1</i>	LTGDETWHHYTPETKEQSKQWIGPGEPTPKKAKTIFSANKVMTVFWDAARGIYID	60
<i>Haematobia irritans</i>	VTIDETWLHHYTPESNRQSAEWTATGEPSPKRGKTQKSADKVMASVFWDAHGIIFID	60
<i>Chrysoperla plorabunda</i>	VTMDETWLHHYTPESKRQSAEWTATGEPSPKRGKTQKSAGKVMASVFWDAHGIIFID	60
<i>Sphyracephala europaea</i>	VAMDETWLHHYTPESNRQSAEWTAYDEPSPKRGKTQQSVGKVMASVFWDAHGIIFID	60
	:: ****:*****::** :* . .*:**:* ** ..*****:***:***:***:	
<i>Agmar1</i>	YLAAGNTITGQYYADLLQRLSHEVKIKRPHLAKKKILFHHDNAPPH-	103
<i>Haematobia irritans</i>	YLEKGKTINSDDYIMALLERLKEVIAAKRPHMKKKKVSFHHDIAPCCHK	103
<i>Chrysoperla plorabunda</i>	YLEKGKTINSDDYIMALLERLKEVIAAKRPHMKKKKVLFHQDNAPCHK	103
<i>Sphyracephala europaea</i>	YLQKGQTINSDDYIALLERLKDEIAAKRPHLKKKKVLFHQDNAPCHK	103
	** ***:**..:** ***:**.*: *****:***:***:***:*** ** *	

B

<i>Agmar2</i>	TGDEKWVYNNVNRKRSWNKKDEPAQTISKADIHKKNVILSVWVDFKGVFFELLF	60
<i>Delphinia picta</i>	TGDEKWIVYNNVNRKRSWSKHGEPAQTTSKADIHQKKVMLSVWVWVDFKGVVY-FELLF	60
<i>Chymomyza amoena</i>	TGDEKWIVYNNVNRKRSWSKHGEPAQTTSKADIHQKKVMLSVWVWVDFKGVVY-FELLF	60
<i>Apis cerana</i>	TGDEKWIVYDNIKR---ELNEPAQRTSKNNIHKK-VMLSVWVDFKGVVY-FELLF	60
	*****:***:***:* : .***** ** :***:* *:*****:***:***: *****	
<i>Agmar2</i>	DNTTINSEVYCHQLDKLNDALQQRSELINRRGVVFHQDNAPPH	92
<i>Delphinia picta</i>	RNQTINSDVYCQQLDKLNAAIKEKRPPELINRKGVIHQDNARPH	92
<i>Chymomyza amoena</i>	RNQTINSDVYCQQLDKLNAAINEKRPPELINRKGVIHQDNARPH	92
<i>Apis cerana</i>	NNCTINSEVYCNQLDKLNNSIKQKRPPELINRKGVIHQDNARPH	92
	* *****:***** ::::***:*****:***:***:*** *	

Fig. 2. Multiple alignment of the deduced amino acid sequences of the soybean aphid *mariner*-like elements (MLEs) with other insect *mariner* transposase sequences. (a) Alignment of *Aphis glycines* MLE transposase sequence (*Agmar1*) with the stalk-eyed fly, *Sphyracephala europaea* (**ABI14926**); common green lacewing, *Chrysoperla plorabunda* (**AAA28265**); and the horn fly, *Haematobia irritans* (**AAA29203**). (b) Alignment of *A. glycines* MLE transposase sequence *Agmar2* with MLE transposase sequences from the Asiatic honey bee, *Aphis cerana* (**BAB86288**); the pomace fly, *Chymomyza amoena* (*Camar1*, **AAO12862**); and the picture-winged fly, *Delphinia picta* (**AAC28155**). In both panels, identical residues among all taxa are indicated by '*', conserved substitutions by ':' and semi-conserved substitutions by '.' symbols.

Southern blot analysis

The soybean aphid *mariner* element *Agmar1* was used as the probe in Southern blot analysis. No *EcoRI* and *XbaI* restriction sites were identified within the nucleotide sequence obtained for *Agmar1*. Five micrograms (representing ~50 aphid individuals) of genomic DNA from both biotypes of *A. glycines* ('OH' and 'IL') was digested individually by *EcoRI* restriction enzyme to generate DNA fragments with intact *Agmar1* sequences. Agarose gel electrophoresis of the DNA fragments was conducted in a 0.8% agarose gel using 1X TAE buffer (0.04 mol l⁻¹ Tris-acetate, 1 mmol l⁻¹ EDTA, pH 8.0). DNA-negative lanes were included in the gel. The gel was treated and the DNA was transferred to Hybond N membrane (Amersham Pharmacia Biotech, Piscataway, NJ, USA) by standard blotting procedure (Sambrook & Russell, 2001). Hybridization was done at 65°C, and probes used were labeled using [α -32P]dCTP (Amersham Pharmacia Biotech) by oligonucleotide random priming (Feinberg & Vogelstein, 1983).

Copy number estimation of *Agmar1* using quantitative PCR

Quantitative PCR has been previously reported in assessing the copy number of a gene (Applied Biosystems, Inc., Foster, CA, USA: Bubner & Baldwin, 2004; Moran et al., 2005; Solomon et al., 2008). The method involves relative

quantification of the gene of interest vs. a reference gene, which is known to be present as a single copy in a haploid genome.

Approximately 20 mg of adult and nymph 'IL' aphids was used for DNA extraction following the Omega Bio-tek protocol described earlier. Two extractions were performed to represent three biological replicates in quantitative PCR (qPCR).

Primers for *Agmar1* (*Agmar1*F2 5'-AGCGTTGTCTGATGAAACA-3' and *Agmar1* R2 5'-ACTACACTCCTGAAACCAAAG-3') and for the reference gene *RPL7* (*RPL7*F2 5'-CACCACGAATACGCATGATGAATG-3' and *RPL7* R2 5'-GCCGAAGCCTATGTAAAGGAGTTC-3') were designed using Beacon Primers 7.5 (Premier Biosoft International, Palo Alto, CA, USA) based on partial genomic DNA and cDNA sequences of *A. glycines*, respectively.

The relative quantification method of CFX 96/C1000 BioRad Real Time System Software was used to approximate the copy number of *Agmar1* in *A. glycines* genome. A standard curve, in which the reference gene ribosomal protein L7 (*RPL7*) is present in 100,000, 10,000, 1000 and 100 copies, was created following steps listed in an Applied Biosystems Inc. online manual. *RPL7* is believed to be present as a single copy gene in the pea aphid (*Acyrtosiphon pisum*) genome and other genomes (Nikoh et al., 2010). Since *A. glycines* genome size is still unknown, the average size (ca. 500 Mb) of *Aphis* sp. (Finston et al., 1995) and *A. pisum* (TIAGC, 2010) was

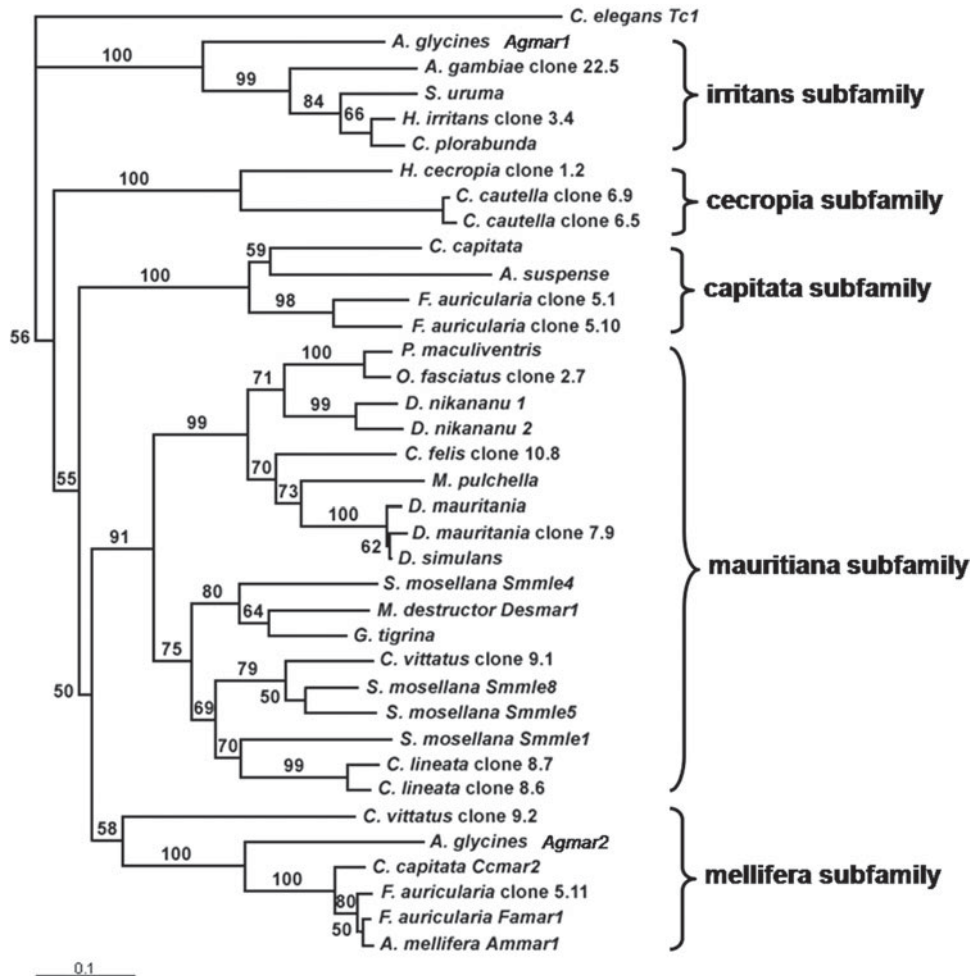


Fig. 3. Phylogenetic analysis showing the relationship of the soybean aphid *mariner*-like elements, *Agmar1* and *Agmar2*, with transposase sequences from the cecropia, irritans, mauritiana, capitata and mellifera subfamilies of *mariner* elements. The phylogram was constructed by the distance/neighbor-joining method. Numbers at the branches correspond to the percentage of 1000 bootstrap replications >50% supporting each branch. Taxa and accession numbers are as follows: *Aphis glycines*, GQ231493 (*Agmar1*) and GQ231494 (*Agmar2*); *Anopheles gambiae*, S36994 (Clone 22.5); *Stomoxys uruma*, U91384; *Haematobia irritans*, S37004 (clone 3.4); *Chrysoperla plorabunda*, AAC46947; *Hyalophora cecropia*, S36996 (clone 1.2); *Cadra cautella*, S37011 (clone 6.9), S37009 (clone 6.5); *Ceratitis capitata*, U04467, AAO12864 (*Ccmar2*); *Anastrepha suspensa*, U04466; *Forficula auricularia*, L10473 (clone 5.1), L10474 (clone 5.10), S3700 (clone 5.11), AAO12863 (*Famar1*); *Podisus maculiventris*, AAC47237; *Oncopeltus fasciatus*, S37012 (clone 2.7); *Drosophila nikananu*, AAC2859, AAC2860; *Ctenocephalides felis*, S37006 (clone 10.8); *Mantispia pulchella*, AAC28152; *Drosophila mauritiana*, AAA28701, S37002 (clone 7.9); *Drosophila simulans*, AAC16616; *Sitodiplosis mosellana*, DQ315365 (*Smmle1*), DQ315366 (*Smmle4*), DQ315367 (*Smmle5*), DQ315368 (*Smmle8*); *Mayetiola destructor*, AAA66077 (*Desmar1*); *Girardia tigrina*, AAB61384; *Chrysops vittatus*, S37017 (clone 9.1), S37018 (clone 9.2); *Ctenolepisma lineata* S37016 (clone 8.7), S37015 (clone 8.6); *Apis mellifera*, AAO12861 (*Ammar1*).

used to calculate the amount of DNA needed to obtain the copy numbers for the standard curve. Spectrophotometric scan at 260 nm was used to determine the amount of DNA needed.

PCR reactions were performed in a 20 µl mixture. The mixture of one reaction contain 1X iQ SYBR Supermix (BioRad, Hercules, CA, USA), 5 µM each of the primers, and 5 µl genomic DNA (about 100 ng). Nuclease-free water (Applied Biosystems, Inc.) was added to make the final volume. The thermal cycling conditions were: 95°C for 5 min, followed by 40 cycles of 95°C for 10 s and 60°C for 30 s. A dissociation step was added to the protocol for a check on

cross-contamination. Two experiments with three replications were conducted.

Results and discussion

PCR amplicons in the size range of 300 bp were cloned from the genome of the soybean aphid (*Aphis glycines*) using degenerate primers designed to target conserved regions of *mariner* transposases (Robertson, 1993; Mittapalli *et al.*, 2006). Similar-sized amplicons were also obtained from the orange blossom wheat midge (*Sitodiplosis mosellana*), which represented the positive control in this study (fig. 1a). No

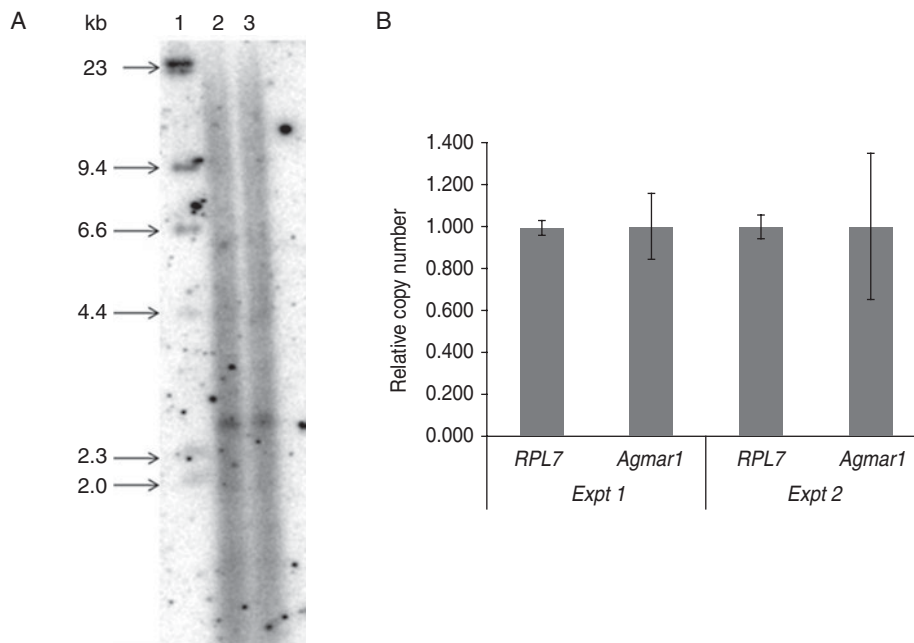


Fig. 4. Copy number estimation of *Agmar1* in *Aphis glycines*. (a) Southern blot analysis of *Aphis glycines* genomic DNA to estimate the copy number of the *mariner*-like element, *Agmar1*. Lanes are as follows: 1, DNA ladder; 2, genomic DNA from 'OH' biotype of *A. glycines* digested with *EcoR1*; 3, genomic DNA from 'IL' biotype of *A. glycines* digested with *EcoR1*. (b) A histogram representing estimated copy number of *Agmar1* relative to a single copy gene, *RPL7*, in two quantitative PCR experiments. Standard error of the mean of three technical replicates is shown.

amplicons were obtained for the negative control, wherein water was used as the template in PCR reactions.

A total of ten PCR clones were sequenced and evaluated as MLE transposases by BLAST searches. Two of the ten *A. glycines* clone sequences translated without in-frame terminations and represented unique sequences that contained conserved residues of *mariner* transposases (fig. 1b). By comparison, three uncharacterized nucleotide sequences categorized as 'Mariner/Tc1' family of transposons exist for the pea aphid, *Acyrthosiphon pisum* Harris (Jurka *et al.*, 2005). Both MLEs from *A. glycines* were designated by the authors as *Agmar1* (clone Sm25-4) and *Agmar2* (clone Sm38) and have the GenBank accessions numbers: **GQ231493** and **GQ231494**, respectively. The other eight cloned sequences were not analyzed further because of multiple stops and in-frame shifts in their deduced protein sequences, suggesting the absence of elements with intact open reading frames and, thus, might not be fully functional (Robertson, 1993). Various BLAST searches, including blastn, blastx and blastp, were performed for *Agmar1* and *Agmar2* to reveal similarities at the nucleotide-nucleotide, nucleotide-protein and (deduced) protein-protein levels. *Agmar1* showed significant identity (E-value=1E-13 and 7E-32; 65% similarity) at the protein level to *mariner* transposase sequences from the common green lacewing (*Chrysoperla plorabunda*), the horn fly (*Haematobia irritans*) and the stable fly (*Stomoxys uruma*). On the other hand, *Agmar2* showed significant (E-value=5E-09 and 2E-20; 70% similarity) at the protein level to *mariner* transposase sequences from the Asiatic honeybee (*Apis cerana*), the picture winged-fly (*Delphinia picta*) and the pomace fly (*Chymomyza amoena*). Alignments of the deduced protein sequences for *Agmar1* and

Agmar2 with the other *mariner* transposase sequences are shown in fig. 2. Thus, the sequence analysis and molecular characterization provide first-hand evidence of the cloned PCR fragments from *A. glycines* genomic DNA are MLE transposase sequences.

Phylogenetic analysis was performed by constructing a neighbor-joining phylogram in accordance with previous studies (Robertson & MacLeod, 1993; Augé-Gouillou *et al.*, 2000). The derived phylogram revealed five clades with good bootstrap support and the topology was in agreement with the expected groupings of the *mariner* transposase sequences into the subfamilies mauritiana, capitata, mellifera, cecropia and irritans (fig. 3). The two MLE transposase sequences from *A. glycines* grouped within the irritans (*Agmar1*) and mellifera (*Agmar2*) subfamilies. Specifically, *Agmar1* grouped with a similar irritans transposase from *Anopheles gambiae*; whereas, *Agmar2* grouped between *Ceratitis capitata* of the mellifera subfamily (fig. 3). These results confirm that the soybean aphid genome hosts at least two different subfamilies of *mariner* transposases. Given the wide distribution of similar elements in an array of arthropods (Robertson, 1993), it is plausible that the *A. glycines* genome hosts other subfamilies of *mariner* transposases than those reported in this study. This divergence of MLEs within a genome could result due to the loss in transposase-DNA interactions, i.e. minor base changes in the ITRs and coding regions of the transposase (Lampe *et al.*, 2001).

Southern blot analysis indicated a low copy number for *Agmar1* within the soybean aphid genomes of both biotypes including 'IL' and 'OH' (fig. 4a). There did not appear to be any differences in copy number between biotypes. Data

obtained via quantitative PCR (qPCR) further supported the Southern analysis, which also revealed a low copy number for *Agmar 1* (fig. 4b). *Agmar 2* was also estimated via qPCR to be low copy number in the genome of *A. glycines* (data not shown).

In insects, the copy numbers of MLEs per genome can be categorized into low, moderate, high and very high (Robertson & Lampe, 1995b). For example, 2–3 copies of MLEs have been estimated within the genomes of *Drosophila ananassae* (Robertson & Lampe, 1995a) and *D. sechellia* (Capy et al., 1992); 20–30 in *Anopheles gambiae* (Robertson & Lampe, 1995b), *D. mauritiana* (Maruyama & Hartl, 1991), *Mayetiola destructor* (Shukle & Russell, 1995), *Sitodiplosis mosellana* (Mittapalli et al., 2006); 1000 in the moth *Hyalophora cecropia* (Lidholm et al., 1991; Robertson, 1993) and up to 17,000 copies in the horn fly, *H. irritans* (Robertson & Lampe, 1995a). Although the genome size of *A. glycines* is likely close to the size of other *Aphis* sp., including *A. pisum* (Finston et al., 1995; TIAGC, 2010), to state an exact copy number for the two retrieved MLEs within the *A. glycines* genome is still uncertain.

Although the function, evolution and transmission of MLEs is not fully known, it is thought that a single copy of these elements initially invades the genome of an organism, perhaps within a virus (Hartl et al., 1997), duplicates through transposition and finally becomes abundant in the population. The most common mechanism for such elements that transpose requiring a DNA intermediate has been proposed to be via horizontal transfer within members of the same order (Maruyama & Hartl, 1991; Simmons, 1992) and/or between members belonging to different insect orders (Robertson & Lampe, 1995b). The low copy numbers of *Agmar1* within the soybean aphid genome suggest the transfer event(s) were fairly recent either being interspecific or intergeneric.

Some intriguing attributes have been assigned to the presence of transposable elements within insect systems. The mobile nature of transposable elements in insects could randomly allow insertions and cause mutations within coding regions of genes (Li et al., 2007). This could lead to the development of resistance against host plant defenses and/or insecticides. Second, MLEs could function as gene transfer systems for many insect species (Maruyama & Hartl, 1991). Hence, these elements can serve as valuable research tools for the isolation and mapping of genes (Cooley et al., 1988), transposon-based sequencing (Strausbaugh et al., 1990) and as vectors to support germline transformation (Lidholm et al., 1993). Although we did not find differences in the genetic makeup of the two soybean aphid *mariners* between biotypes, there could be other MLEs that exist within the genome that may influence their behavior on the host plants. Nevertheless, both *Agmar1* and *Agmar2* clearly represent useful tools to map and isolate genes in the soybean aphid genome. Additionally, these mobile DNA elements within the genome of *A. glycines* might be an important source for biotype evolution and/or genetic diversity within populations.

Further studies need to be performed to learn whether the low copy numbers of MLEs found in the *A. glycines* genome are active or inactive. For this to be achieved, full length sequences need to be obtained, and subsequent *in vitro* genetic analysis for transposition such as the 'hop' assays developed for *Mboumar-9* (Muñoz-Lopez et al., 2008) need to be performed.

Acknowledgements

We thank Dr Esther van der Knaap's laboratory for help with Southern blot analysis. Help with sequencing at the Molecular and Cellular Imaging Center (MCIC, Wooster, OH) is much appreciated. We also thank Dr Curt Hill (University of Illinois) for sending the Illinois biotype aphids. The authors appreciate the comments/suggestions provided by the reviewers. Funding for this project was provided by State and Federal funds appropriated to the Ohio Agricultural Research and Development Center (OARDC, Wooster, OH), The Ohio State University and by the Ohio Soybean Council (OSC 08-2-08).

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410.
- Augé-Gouillou, C., Notareschi-Leroy, H., Abad, P., Periquet, G. & Bigot, Y. (2000) Phylogenetic analysis of the functional domains of *mariner*-like element (MLE) transposases. *Molecular and General Genetics* **264**, 506–513.
- Barry, E.G., Witherspoon, D.J. & Lampe, D.J. (2004) A bacterial genetic screen identifies functional coding sequences of the insect *mariner* transposable element *Famar 1* amplified from the genome of the earwig, *Forficula auricularia*. *Genetics* **166**, 823–833.
- Bryan, G.J., Jacobson, J.W. & Hartl, D.L. (1987) Heritable somatic excision of a *Drosophila* transposon. *Science* **235**, 1636–1638.
- Bubner, B. & Baldwin, I.T. (2004) Use of real time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Reports* **23**, 263–271.
- Capy, P., David, J.R. & Hartl, D.L. (1992) Evolution of the transposable element *mariner* in the *Drosophila melanogaster* species group. *Genetica* **86**, 37–46.
- Cooley, L., Kelley, R. & Spradling, A. (1988) Insertional mutagenesis of the *Drosophila* genome with single P elements. *Science* **239**, 1121–1128.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797.
- FAO Statistical Database (2010) FAOSTAT agriculture data. Available online at <http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor> (accessed 28 November 2010).
- Feinberg, A.P. & Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**, 6–13.
- Finston, T.L., Hebert, P.D.N. & Footitt, R. (1995) Genome size variation in aphids. *Insect Biochemistry and Molecular Biology* **25**, 189–196.
- Hartl, D.L. (1989) Transposable element *mariner* in *Drosophila* species. pp. 531–536 in Berg, D.E. & Howe, M.M. (Eds) *Mobile DNA*. Washington, DC, USA, American Society of Microbiology.
- Hartl, D.L., Lohe, A.R. & Lozovskaya, E.R. (1997) Modern thoughts on an ancient *marinere*: function, evolution, regulation. *Annual Review of Genetics* **31**, 337–358.
- Hill, C.B., Li, Y. & Hartman, G.L. (2004) Resistance to the soybean aphid in soybean germplasm. *Crop Science* **44**, 98–106.
- Hill, C.B., Crull, L., Herman, T.K., Voegtlin, D.J. & Hartman, G.L. (2010) A new soybean aphid (Hemiptera: Aphididae)

- biotype identified. *Journal of Economic Entomology* **103**, 509–515.
- Izsvak, Z., Ivics, Z., Shimoda, N., Mohn, D., Okamoto, H. & Hackett, P.B. (1999) Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification. *Journal of Molecular Evolution* **48**, 13–21.
- Jacobson, J.W., Medhora, M.M. & Hartl, D.L. (1986) Molecular structure of a somatically unstable transposable element in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 8684–8688.
- Jarvik, T. & Lark, K.G. (1998) Characterization of Soyamarl, a *mariner* element in soybean. *Genetics* **149**, 1569–1574.
- Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O. & Walichiewicz, J. (2005) Repbase update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research* **110**, 462–467.
- Kim, K.S., Hill, C.B., Hartman, G.L., Mian, R.M. & Diers, B.W. (2008) Discovery of soybean aphid biotypes. *Crop Science* **48**, 923–928.
- Lampe, D.J., Walden, K.K.O. & Robertson, H.M. (2001) Loss of transposase–DNA interaction may underlie the divergence of *mariner* family transposable elements and the ability of more than one *mariner* to occupy the same genome. *Molecular Biology and Evolution* **18**, 954–961.
- Leroy, H., Leroy, F., Auge-Gouillou, C., Castagnone-Sereno, P. & Vanlerberghe-Masutti, F. (2000) Identification of *mariner*-like elements from the root-knot nematode *Meloidogyne* spp. *Molecular and Biochemical Parasitology* **107**, 181–190.
- Li, X., Schuler, M.A. & Berenbaum, M.R. (2007) Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annual Review of Entomology* **52**, 231–253.
- Lidholm, D.A., Lohe, A.R. & Hartl, D.L. (1993) The transposable element *mariner* mediates germline transformation in *Drosophila melanogaster*. *Genetics* **134**, 859–868.
- Liu, D., Bischerour, J., Siddique, A., Buisine, N., Bigot, Y. & Chalmers, R. (2007) The human SETMAR protein preserves most of the activities of the ancestral *Hsmar1* transposase. *Molecular and Cellular Biology* **27**, 1125–1132.
- Maruyama, K. & Hartl, D.L. (1991) Evolution of the transposable element *mariner* in *Drosophila* species. *Genetics* **128**, 319–329.
- Medhora, M., Maruyama, K. & Hartl, D.L. (1991) Molecular and functional analysis of the *mariner* mutator element *Mos1* in *Drosophila*. *Genetics* **128**, 311–318.
- Mensah, C., DiFonzo, C., Nelson, R.L. & Wang, D. (2005) Resistance to soybean aphid in early maturing soybean germplasm. *Crop Science* **45**, 2228–2233.
- Mian, M.A.R., Hammond, R.B. & Martin, S.K. (2008a) New plant introductions with resistance to the soybean aphid. *Crop Science* **48**, 1055–1061.
- Mian, M.A.R., Kang, S.-T., Beil, S.E. & Hammond, R.B. (2008b) Genetic linkage mapping of the soybean aphid resistance gene in PI 243540. *Theoretical and Applied Genetics* **117**, 955–962.
- Mittapalli, O., Shukle, R.H. & Wise, I.L. (2006) Identification of *mariner*-like elements in *Sitodiplosis mosellana* (Diptera: Cecidomyiidae). *Canadian Entomologist* **138**, 138–146.
- Moran, N.A., Degnan, P.H., Santos, S.R., Dunbar, H.E. & Ochman, H. (2005) The players in a mutualistic symbiosis: Insects, bacteria, viruses and virulence genes. *Proceedings of the National Academy of Sciences* **102**, 16919–16926.
- Muñoz-Lopez, M., Siddique, A., Bischerour, J., Lorite, P., Chalmers, R. & Palomeque, T. (2008) Transposition of *Mboumar-9*: identification of a new naturally active *mariner*-family transposon. *Journal of Molecular Biology* **382**, 567–572.
- Nikoh, N., McCutcheon, J.P., Kudo, T., Miyagishima, S., Moran, N.A. & Nakabochi, A. (2010) Bacterial genes in the aphid genome: absence of functional gene transfer from *Buchnera* to its hosts. *PLoS Genetics* **6**, e1000827.
- Ragsdale, D.W., McCornack, B.P., Venette, R.C., Potter, B.D., MacRae, I.V., Hodgson, E.W., O'Neal, M.E., Johnson, K.D., O'Neil, R.J., DiFonzo, C.D., Hunt, T.E., Glogoza, P.A. & Cullen, E.M. (2007) Economic threshold for soybean aphid (Hemiptera: Aphididae). *Journal of Economic Entomology* **100**, 1258–1267.
- Robertson, H.M. (1993) The *mariner* transposable element is widespread in insects. *Nature* **362**, 241–245.
- Robertson, H.M. & Lampe, D.J. (1995a) Distribution of transposable elements in arthropods. *Annual Review of Entomology* **40**, 333–357.
- Robertson, H.M. & Lampe, D.J. (1995b) Recent horizontal transfer of a *mariner* transposable element among and between Diptera and Neuroptera. *Molecular Biology and Evolution* **12**, 850–862.
- Robertson, H.M. & MacLeod, E.G. (1993) Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. *Insect Molecular Biology* **2**, 125–139.
- Sambrook, J. & Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*. 3rd edn. Cold Spring Harbor, NY, USA, Cold Spring Harbor Laboratory Press.
- Schneider, D., Faure, D., Noirclerc-Savoye, M., Barriere, A.C., Coursange, E. & Blot, M. (2000) A broad-host-range plasmid for isolating mobile genetic elements in gram-negative bacteria. *Plasmid* **44**, 201–207.
- Shukle, R.H. & Russell, V.W. (1995) *Mariner* transposase-like sequences from the Hessian fly, *Mayetiola destructor*. *Journal of Heredity* **86**, 364–368.
- Simmons, G.M. (1992) Horizontal transfer of hobo transposable elements within the *Drosophila melanogaster* species complex, evidence from DNA sequencing. *Molecular Biology and Evolution* **9**, 1050–1060.
- Solomon, P.S., Ipcho, S.V.S., Hane, J.K., Tan, K.C. & Oliver, R.P. (2008) A quantitative PCR approach to determine gene copy number. *Fungal Genetics Reports* **55**, 5–8.
- Strausbaugh, L.D., Bourke, M.T., Sommer, M.T., Coon, M.E. & Berg, C.M. (1990) Probe mapping to facilitate transposon-based DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 6213–6217.
- Swofford, D.L. (2002) PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4. Sunderland, MA, USA, Sinauer Associates.
- TIAGC (The International Aphid Genomics Consortium) (2010) Genome Sequence of the Pea Aphid *Acyrtosiphon pisum*. *PLoS Biology* **8**, e1000313.
- Thompson, J.D., Gibson, T.J., Plewniak, F. & Jeanmougin, F. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**, 4876–4882.
- Venette, R.C. & Ragsdale, D.W. (2004) Assessing the invasion by soybean aphid (Homoptera: Aphididae): Where will it end? *Annals of the Entomological Society of America* **97**, 219–226.
- Wu, Z., Schenk-Hamlin, D., Zhan, W., Ragsdale, D.W. & Heimpel, G.E. (2004) The soybean aphid in China: an historical review. *Annals of the Entomological Society of America* **97**, 209–218.
- Zhang, G.R., Gu, C.H. & Wang, D.C. (2009) Molecular mapping of soybean aphid resistance genes in PI 567541B. *Theoretical and Applied Genetics* **118**, 473–482.